两位玩家，A、B（或者更多，简单来说暂定就两位）玩一个大富翁游戏。

游戏地图：

一个圆环，共计15个（暂定）格子。

游戏规则：

玩家初始拥有一定现金，轮流掷骰子（1~6），玩家到达空地可以选择造房子或啥都不做，到达其他玩家的房子必须根据房子等级支付一定费用。

每一轮会固定扣除玩家一定的金额，玩家破产时游戏结束。

附加游戏规则（可选）：

1. 房子为玩家的不动产（金额低于购入时的价格），没有现金时可以变卖房子抵债。
2. 玩家到自己的房子可以选择花钱升级房子。
3. 有不同类型的格子。例如普通空地，特殊空地，商店，奖励格子（走上去拿钱），银行（存款），彩票，随机事件等。
4. 玩家建造的建筑类型发生变化，普通空地造普通房子，特殊空地可以造酒店，科技站等。
5. 玩家可以使用不同种类的道具。例如遥控骰子，路障等，有上限。
6. 不同类型的玩家有不同的技能。

暂定解决思路：

用Markov decision process (MDP)方法建模。估计当前状态下所有决策后的状态的期望价值，选取最高的那个对应的决策作为玩家的行为。

方法一：值迭代（状态空间大，定义终结态）。

方法二：策略迭代（状态空间大，定义终结态）。

方法三：决策树，计算未来一定轮数内的所有情况，选择综合期望价值最高的决策。

方法四：强化学习，定义状态特征（例如玩家位置状态，各个玩家的房子数，当前金额，等等）训练评估函数的参数（即时反馈和输赢反馈），计算状态价值。例：Value = 10\*玩家房子数-5\*对方房子数+1.5\*当前金额。

方法五：神经网络，将状态编码并输入网络，输出预测的价值。